

## **Ferret: a user-friendly tool to extract data from the 1000 Genomes Project**

A. Taverner, O. Belkacem, E. Lam, N. Vince, P.-A. Gourraud, C. Winkler, M.  
Servières, Sophie Limou

### **► To cite this version:**

A. Taverner, O. Belkacem, E. Lam, N. Vince, P.-A. Gourraud, et al.. Ferret: a user-friendly tool to extract data from the 1000 Genomes Project. Assises de Génétique Humaine et Médicale, Jan 2018, Nantes, France. inserm-02163244

**HAL Id: inserm-02163244**

**<https://www.hal.inserm.fr/inserm-02163244>**

Submitted on 24 Jun 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Ferret: a user-friendly tool to extract data from the 1000 Genomes Project



**Contact:**  
sophie.limou@univ-nantes.fr  
ferret@nih.gov

A. Taverner<sup>1,2</sup>, O. Belkacem<sup>3</sup>, E. Lam<sup>3</sup>, N. Vince<sup>4,5</sup>, PA. Gourraud<sup>4,5</sup>,  
CA. Winkler<sup>1</sup>, M. Servières<sup>3</sup>, S. Limou<sup>1,3-5</sup>

<sup>1</sup>Basic Research Laboratory, NCI/NIH, Leidos Biomedical Research Inc., Frederick MD, USA

<sup>2</sup>Quantitative and Computational Biology program, Princeton University, Princeton NJ, USA

<sup>3</sup>Ecole Centrale de Nantes, Nantes, France

<sup>4</sup>Centre de Recherche en Transplantation et Immunologie (CRTI) UMR 1064, INSERM, Nantes, France

<sup>5</sup>Institut de Transplantation Urologie Néphrologie (ITUN), CHU Nantes, Nantes, France

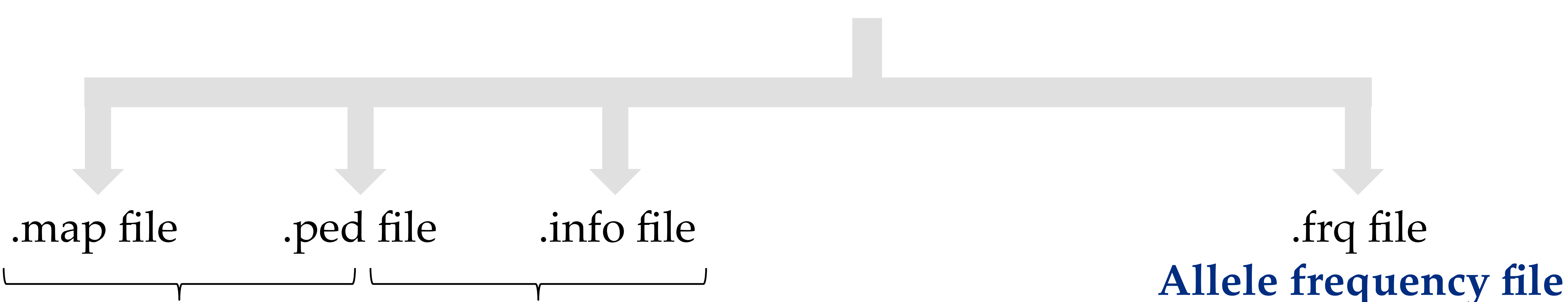
## Additional settings:

- Genome version
- Output format
- Frequency filter

## Input flexibility:

- Locus position
- Gene
- Variant ID

Retrieve 1000 Genomes Project data  
(and optionally Exome Sequencing Project –ESP– data)



### PLINK files

can be merged with user's dataset, etc.

### HaploView files

data visualization, LD pattern, haplotypes, tagSNP design, etc.

CHR	VARIANT	POS	A1	A2	NB_1KG_	1KG_A1	ESP6500_EA	ESP6500_AA
					CHR	_FREQ	_A1_FREQ	_A1_FREQ
6	rs1065711	31236853	G	A	1006	0.1213	.	.
6	rs1049853	31236900	G	A	1006	0.8867	0.8726	0.9209
6	indel_rs41548117_TC/T	31237100	TC	T	1006	0.9553	0.9594	0.9183
6	rs41542414	31237766	A	T	1006	1	0.9995	0.9897
6	rs9264623	31237950	T	C	1006	0.2773	.	.
6	indel_rs9281300_C/CA	31239170	C	CA	1006	0.4195	0.535	0.4125

## Unique advantages of Ferret<sup>#</sup>

- User-friendly interface
- Accepts input query as locus, gene(s), or variant(s)
- Handles SNPs and indels
- Outputs suitable for well-known pre-existing tools
- Computes allelic frequency for SNPs, indels and CNVs
- Retrieves allelic frequency from ESP
- **Dec. 2017 update:** sync with HaploView + gene borders

## Perspectives (June 2018):

- Accelerate runtime speed
- Grant access to HLA alleles<sup>§</sup>
- Functional annotations

Publicly available at <http://limousophie35.github.io/Ferret/>